

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/766,366

DATE: 04/26/2001

TIME: 17:25:55

Input Set : N:\Crf3\RULE60\09766366.txt

Output Set: N:\CRF3\04262001\I766366.raw

SEQUENCE LISTING

ENTERED

C--> 3 (1) GENERAL INFORMATION:
 5 (i) APPLICANT: Hillman, Jennifer L.
 6 Shah, Purvi
 7 Corley, Neil C.
 C--> 9 (ii) TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
 11 (iii) NUMBER OF SEQUENCES: 4
 13 (iv) CORRESPONDENCE ADDRESS:
 14 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 15 (B) STREET: 3174 Porter Drive
 16 (C) CITY: Palo Alto
 17 (D) STATE: CA
 18 (E) COUNTRY: USA
 19 (F) ZIP: 94304
 21 (v) COMPUTER READABLE FORM:
 22 (A) MEDIUM TYPE: Diskette
 23 (B) COMPUTER: IBM Compatible
 24 (C) OPERATING SYSTEM: DOS
 25 (D) SOFTWARE: FastSEQ for Windows Version 2.0
 27 (vi) CURRENT APPLICATION DATA:
 C--> 28 (A) APPLICATION NUMBER: US/09/766,366
 C--> 29 (B) FILING DATE: 18-Jan-2001
 31 (vii) PRIOR APPLICATION DATA:
 32 (A) APPLICATION NUMBER: 08/872,784
 33 (B) FILING DATE:
 35 (viii) ATTORNEY/AGENT INFORMATION:
 36 (A) NAME: Billings, Lucy J.
 37 (B) REGISTRATION NUMBER: 36,749
 38 (C) REFERENCE/DOCKET NUMBER: PF-0293 US
 40 (ix) TELECOMMUNICATION INFORMATION:
 41 (A) TELEPHONE: 415-855-0555
 42 (B) TELEFAX: 415-845-4166
 44 (2) INFORMATION FOR SEQ ID NO: 1:
 46 (i) SEQUENCE CHARACTERISTICS:
 47 (A) LENGTH: 311 amino acids
 48 (B) TYPE: amino acid
 49 (C) STRANDEDNESS: single
 50 (D) TOPOLOGY: linear
 52 (vii) IMMEDIATE SOURCE:
 53 (A) LIBRARY: BRAINOT09
 54 (B) CLONE: 2150905
 56 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 58 Met Gly Arg Ala Val Ala Thr Ala Ala Leu Pro Pro Gly Asp Leu Arg
 59 1 5 10 15
 60 Ser Val Leu Val Thr Thr Val Leu Asn Leu Glu Pro Leu Asp Glu Asp
 61 20 25 30
 62 Leu Phe Arg Gly Arg His Tyr Trp Val Pro Ala Lys Arg Leu Phe Gly

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/766,366

DATE: 04/26/2001

TIME: 17:25:55

Input Set : N:\Crf3\RULE60\09766366.txt

Output Set: N:\CRF3\04262001\I766366.raw

```

63          35          40          45
64 Gly Gln Ile Val Gly Gln Ala Leu Val Ala Ala Ala Lys Ser Val Ser
65          50          55          60
66 Glu Asp Val His Val His Ser Leu His Cys Tyr Phe Val Arg Ala Gly
67          65          70          75          80
68 Asp Pro Lys Leu Pro Val Leu Tyr Gln Val Glu Arg Thr Arg Thr Gly
69          85          90          95
70 Ser Ser Phe Ser Val Arg Ser Val Lys Ala Val Gln His Gly Lys Pro
71          100         105         110
72 Ile Phe Ile Cys Gln Ala Ser Phe Gln Gln Ala Gln Pro Ser Pro Met
73          115         120         125
74 Gln His Gln Phe Ser Met Pro Thr Val Pro Pro Pro Glu Glu Leu Leu
75          130         135         140
76 Asp Cys Glu Thr Leu Ile Asp Gln Tyr Leu Arg Asp Pro Asn Leu Gln
77          145         150         155         160
78 Lys Arg Tyr Pro Leu Ala Leu Asn Arg Ile Ala Ala Gln Glu Val Pro
79          165         170         175
80 Ile Glu Ile Lys Pro Val Asn Pro Ser Pro Leu Ser Gln Leu Gln Arg
81          180         185         190
82 Met Glu Pro Lys Gln Met Phe Trp Val Arg Ala Arg Gly Tyr Ile Gly
83          195         200         205
84 Glu Gly Asp Met Lys Met His Cys Cys Val Ala Ala Tyr Ile Ser Asp
85          210         215         220
86 Tyr Ala Phe Leu Gly Thr Ala Leu Leu Pro His Gln Trp Gln His Lys
87          225         230         235         240
88 Val His Phe Met Val Ser Leu Asp His Ser Met Trp Phe His Ala Pro
89          245         250         255
90 Phe Arg Ala Asp His Trp Met Leu Tyr Glu Cys Glu Ser Pro Trp Ala
91          260         265         270
92 Gly Gly Ser Arg Gly Leu Val His Gly Arg Leu Trp Arg Gln Asp Gly
93          275         280         285
94 Val Leu Ala Val Thr Cys Ala Gln Glu Gly Val Ile Arg Val Lys Pro
95          290         295         300
96 Gln Val Ser Glu Ser Lys Leu
97          305         310
99 (2) INFORMATION FOR SEQ ID NO: 2:
101   (i) SEQUENCE CHARACTERISTICS:
102       (A) LENGTH: 1098 base pairs
103       (B) TYPE: nucleic acid
104       (C) STRANDEDNESS: single
105       (D) TOPOLOGY: linear
107   (vii) IMMEDIATE SOURCE:
108       (A) LIBRARY: BRAINOT09
109       (B) CLONE: 2150905
111   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
113 CAGCATTGAA CTAGATGTCG TCCCCGCAGG CCCCAGAAGA TGGGCAGGGC TGTGGCGACC      60
114 GCGGCGCTTC CCCCTGGGGA CCTCCGTAGC GTCTTGGTCA CGACCGTGCT CAACCTCGAG      120
115 CCGCTGGACG AGGATCTCTT CAGAGGAAGG CATTACTGGG TACCGGCCAA GAGGCTGTTT      180
116 GGTGGTCAGA TCGTGGGCCA GGCCCTGGTG GCTGCAGCCA AGTCTGTGAG TGAAGACGTC      240

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/766,366

DATE: 04/26/2001

TIME: 17:25:55

Input Set : N:\Crf3\RULE60\09766366.txt

Output Set: N:\CRF3\04262001\I766366.raw

```

117 CACGTGCACT CCCTGCACTG CTACTTTGTT CGGGCAGGGG ACCCGAAGCT GCCAGTACTG 300
118 TACCAAGTGG AGCGGACACG AACAGGGTCG AGCTTCTCGG TGCCTCTGT GAAGGCCGTG 360
119 CAACATGGGA AGCCCATCTT CATCTGCCAG GCCTCCTTCC AGCAGGCCCA GCCCAGCCCC 420
120 ATGCAGCACC AGTTCTCCAT GCCCACTGTG CCACCACCAG AAGAGCTGCT TGACTGTGAG 480
121 ACCCTCATTT ACCAGTATTT AAGGGACCCT AACCTCCAAA AGAGGTACCC ATTGGCGCTC 540
122 AACCGAATTG CTGCTCAGGA GGTCCCCATT GAGATCAAGC CAGTAAACCC ATCCCCCCTG 600
123 AGCCAGCTGC AGAGAATGGA GCCCAAACAG ATGTTCTGGG TGCGAGCCCG GGGCTATATT 660
124 GGCGAGGGCG ACATGAAGAT GCACTGCTGC GTGGCCGCCT ATATCTCCGA CTATGCCTTC 720
125 TTGGGCACTG CACTGCTGCC TCACCAGTGG CAGCACAAGG TGCACTTCAT GGTCTCACTG 780
126 GACCATTCCA TGTGGTTCCA CGCCCCCTTC CGAGCTGACC ACTGGATGCT CTATGAATGC 840
127 GAGAGCCCCT GGGCCGGTGG CTCTCGGGGG CTGTTCCATG GGCGGCTGTG GCGTCAGGAT 900
128 GGAGTCCTAG CTGTGACCTG TGCCCAGGAG GCGGTGATCC GAGTGAAGCC CCAGGTCTCA 960
129 GAGAGCAAGC TGTAGCCAGA GGTACCAGCT TCGCCTGGGG CTTCAAGAAC CTCCCATCTA 1020
130 TCCCCATTCC TGAGACAGGA GTTACAGTCC CTTTGGGCC TCACATCCAA TAAAGAGACT 1080
131 GATACCACTG GAAAAAAA
1098

```

133 (2) INFORMATION FOR SEQ ID NO: 3:

135 (i) SEQUENCE CHARACTERISTICS:

136 (A) LENGTH: 286 amino acids

137 (B) TYPE: amino acid

138 (C) STRANDEDNESS: single

139 (D) TOPOLOGY: linear

141 (vii) IMMEDIATE SOURCE:

142 (A) LIBRARY: GenBank

143 (B) CLONE: 147932

145 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

147 Met Ser Gln Ala Leu Lys Asn Leu Leu Thr Leu Leu Asn Leu Glu Lys
148 1 5 10 15
149 Ile Glu Glu Gly Leu Phe Arg Gly Gln Ser Glu Asp Leu Gly Leu Arg
150 20 25 30
151 Gln Val Phe Gly Gly Gln Val Val Gly Gln Ala Leu Tyr Ala Ala Lys
152 35 40 45
153 Glu Thr Val Pro Glu Glu Arg Leu Val His Ser Phe His Ser Tyr Phe
154 50 55 60
155 Leu Arg Pro Gly Asp Ser Lys Lys Pro Ile Ile Tyr Asp Val Glu Thr
156 65 70 75 80
157 Leu Arg Asp Gly Asn Ser Phe Ser Ala Arg Arg Val Ala Ala Ile Gln
158 85 90 95
159 Asn Gly Lys Pro Ile Phe Tyr Met Thr Ala Ser Phe Gln Ala Pro Glu
160 100 105 110
161 Ala Gly Phe Glu His Gln Lys Thr Met Pro Ser Ala Pro Ala Pro Asp
162 115 120 125
163 Gly Leu Pro Ser Glu Thr Gln Ile Ala Gln Ser Leu Ala His Leu Leu
164 130 135 140
165 Pro Pro Val Leu Lys Asp Lys Phe Ile Cys Asp Arg Pro Leu Glu Val
166 145 150 155 160
167 Arg Pro Val Glu Phe His Asn Pro Leu Lys Gly His Val Ala Glu Pro
168 165 170 175
169 His Arg Gln Val Trp Ile Arg Ala Asn Gly Ser Val Pro Asp Asp Leu
170 180 185 190

```

RAW SEQUENCE LISTING

DATE: 04/26/2001

PATENT APPLICATION: US/09/766,366

TIME: 17:25:55

Input Set : N:\Crif3\RULE60\09766366.txt

Output Set: N:\CRF3\04262001\I766366.raw

```

171 Arg Val His Gln Tyr Leu Leu Gly Tyr Ala Ser Asp Leu Asn Phe Leu
172      195                        200                        205
173 Pro Val Ala Leu Gln Pro His Gly Ile Gly Phe Leu Glu Pro Gly Ile
174      210                        215                        220
175 Gln Ile Ala Thr Ile Asp His Ser Met Trp Phe His Arg Pro Phe Asn
176      225                        230                        235                        240
177 Leu Asn Glu Trp Leu Tyr Ser Val Glu Ser Thr Ser Ala Ser Ser
178      245                        250                        255
179 Ala Arg Gly Phe Val Arg Gly Glu Phe Tyr Thr Gln Asp Gly Val Leu
180      260                        265                        270
181 Val Ala Ser Thr Val Gln Glu Gly Val Met Arg Asn His Asn
182      275                        280                        285
184 (2) INFORMATION FOR SEQ ID NO: 4:
186     (i) SEQUENCE CHARACTERISTICS:
187         (A) LENGTH: 349 amino acids
188         (B) TYPE: amino acid
189         (C) STRANDEDNESS: single
190         (D) TOPOLOGY: linear
192     (vii) IMMEDIATE SOURCE:
193         (A) LIBRARY: GenBank
194         (B) CLONE: 854594
196     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
198 Met Ser Ala Ser Lys Met Ala Met Ser Asn Leu Glu Lys Ile Leu Glu
199   1      5      10      15
200 Leu Val Pro Leu Ser Pro Thr Ser Phe Val Thr Lys Tyr Leu Pro Ala
201      20      25      30
202 Ala Pro Val Gly Ser Lys Gly Thr Phe Gly Gly Thr Leu Val Ser Gln
203      35      40      45
204 Ser Leu Leu Ala Ser Leu His Thr Val Pro Leu Asn Phe Phe Pro Thr
205      50      55      60
206 Ser Leu His Ser Tyr Phe Ile Lys Gly Gly Asp Pro Arg Thr Lys Ile
207      65      70      75      80
208 Thr Tyr His Val Gln Asn Leu Arg Asn Gly Arg Asn Phe Ile His Lys
209      85      90      95
210 Gln Val Ser Ala Tyr Gln His Asp Lys Leu Ile Phe Thr Ser Met Ile
211      100     105     110
212 Leu Phe Ala Val Gln Arg Ser Lys Glu His Asp Ser Leu Gln His Trp
213      115     120     125
214 Glu Thr Ile Pro Gly Leu Gln Gly Lys Gln Pro Asp Pro His Arg Tyr
215      130     135     140
216 Glu Glu Ala Thr Ser Leu Phe Gln Lys Glu Val Leu Asp Pro Gln Lys
217      145     150     155     160
218 Leu Ser Arg Tyr Ala Ser Leu Ser Asp Arg Phe Gln Asp Ala Thr Ser
219      165     170     175
220 Met Ser Lys Tyr Val Asp Ala Phe Gln Tyr Gly Val Met Glu Tyr Gln
221      180     185     190
222 Phe Pro Lys Asp Met Phe Tyr Ser Ala Arg His Thr Asp Glu Leu Asp
223      195     200     205
224 Tyr Phe Val Lys Val Arg Pro Pro Ile Thr Thr Val Glu His Ala Gly

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/766,366

DATE: 04/26/2001

TIME: 17:25:55

Input Set : N:\Crif3\RULE60\09766366.txt

Output Set: N:\CRF3\04262001\I766366.raw

```

225      210      215      220
226 Asp Glu Ser Ser Leu His Lys His His Pro Tyr Arg Ile Pro Lys Ser
227 225      230      235      240
228 Ile Thr Pro Glu Asn Asp Ala Arg Tyr Asn Tyr Val Ala Phe Ala Tyr
229      245      250      255
230 Leu Ser Asp Ser Tyr Leu Leu Leu Thr Ile Pro Tyr Phe His Asn Leu
231      260      265      270
232 Pro Leu Tyr Cys His Ser Phe Ser Val Ser Leu Asp His Thr Ile Tyr
233      275      280      285
234 Phe His Gln Leu Pro His Val Asn Asn Trp Ile Tyr Leu Lys Ile Ser
235      290      295      300
236 Asn Pro Arg Ser His Trp Asp Lys His Leu Val Gln Gly Lys Tyr Phe
237 305      310      315      320
238 Asp Thr Gln Ser Gly Arg Ile Met Ala Ser Val Ser Gln Glu Gly Tyr
239      325      330      335
240 Val Val Tyr Gly Ser Glu Arg Asp Ile Arg Ala Lys Phe
241      340      345

```

VERIFICATION SUMMARY

DATE: 04/26/2001

PATENT APPLICATION: US/09/766,366

TIME: 17:25:56

Input Set : N:\CrF3\RULE60\09766366.txt

Output Set: N:\CRF3\04262001\I766366.raw

L:3 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:9 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]